ID: 816

Key limitations to aquatic eDNA metabarcoding: a cautionary case-study from a diverse public aquarium

Kevin C Morey, Timothy J Bartley, Robert H Hanner

Biodiversity Institute of Ontario, University of Guelph, Guelph, N1G 2W1, Ontario, Canada

ABSTRACT

Environmental DNA (eDNA) and DNA metabarcoding techniques have been widely touted as powerful new tools for monitoring biodiversity in aquatic ecosystems. However, these techniques are still in their infancy and require thorough validation because there are still several key uncertainties surrounding eDNA metabarcoding. These uncertainties include both methodological and analytical limitations that must be addressed before the wide adoption of eDNA metabarcoding for biodiversity monitoring. In this study, we assess both the ability for eDNA metabarcoding to capture biodiversity in a highly diverse closed system at the Ripley's Aquarium of Canada in Toronto, Ontario as well as address a number of knowledge gaps pertaining to eDNA metabarcoding to open a discussion on current issues limiting this tool. Results: This study found that eDNA metabarcoding recovered 62 of 107 (58%) target species and 30 of 44 (68%) target genera from a closed system when using a multi-marker (*COI, 16S, 12S*) approach. Additionally, individual markers showed great disparity in off-target identification noise with *COI* producing the greatest proportion of noise (95% of OTUs). This case study represents the first to highlight key uncertainties and current challenges for eDNA metabarcoding as a biodiversity monitoring tool in highly diverse closed aquatic ecosystems. We identify several outstanding issues with eDNA metabarcoding relating to contamination, sampling methodology, study design, statistical and bioinformatic analyses, and a lack of standardized protocols. These issues raise concerns for the reliability of eDNA metabarcoding when applied to studying complex and highly diverse natural systems. These concerns are reminiscent of those identified previously for DNA barcoding and ancient DNA work. We conclude that the key facets of eDNA metabarcoding methodology that we identify here require further focus before eDNA metabarcoding can be broadly applied in aquatic biodiversity monitoring.

INTRODUCTION

- Monitoring and surveying complex aquatic ecosystems with high levels of diversity is a significant challenge due to the amount of expertise, time, effort, and money required to do it successfully
- Effectiveness of eDNA metabarcoding has been evaluated by a limited number of studies in species-poor artificial ecosystems[1]
- Using markers for *16S*, *12S*, and *COI* we assessed the utility of eDNA metabarcoding by sampling the Rainbow Reef at Ripley's Aquarium of Canada – a diverse and complex system containing 107 species of marine fishes.

METHODS

- Three 1-L water samples were collected from the Rainbow Reef tank and one 1-L water sample was collected from pre-tank Loop Line water system on-site.
- Library preparation was conducted for previously published metabarcoding primers for *16S*[2], *12S*[3], and *COI*[4].
- NGS conducted using Illumina MiSeq and sequences processed bioinformatically using uSEARCH pipeline

KEY LIMITATIONS FOR AQUATIC eDNA METABARCODING

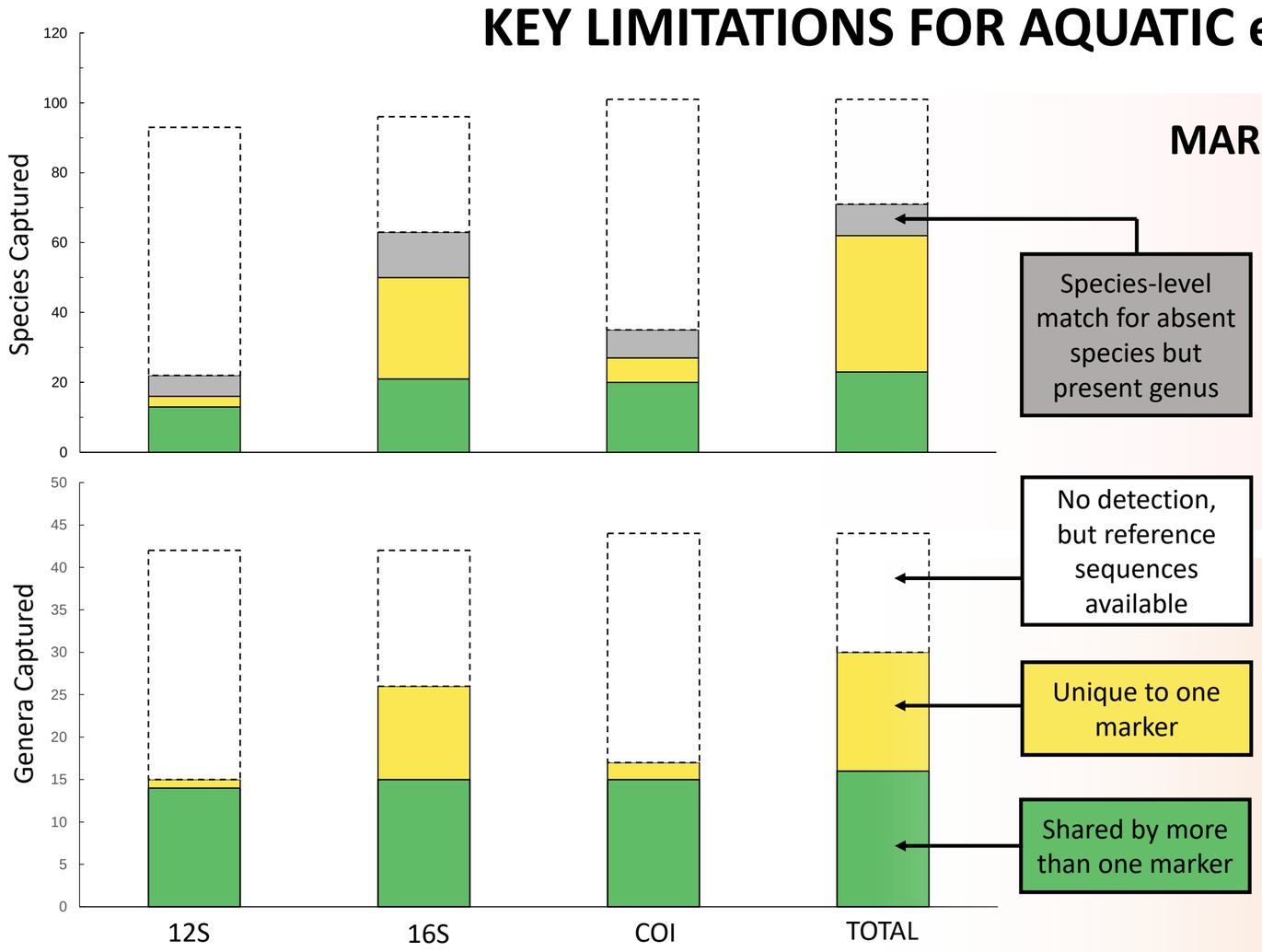

Figure 1. Species-(top) and Genus-Level (bottom) Identification Matches. The selected 16S marker was found to produce the most shared and unique identifications and had the least number of missed identifications, however maximum detection success only occurred when combining all markers.

MARKER SELECTION AND USE OF MULTIPLE MARKERS

- Found maximum total detection and unique detection of species and genera using data from all combined markers (Figure 1).
- Use of individual markers limits detection potential
  - Need to be able to account for marker-specific biases[5,6]
- Will increase cost and data processing time, but would produce more robust data

CONTROLLING FOR CONTAMINATION

- Found marker-specific differences in off-target/contamination sequences detected (Figure 2).
- Management of false positive and false negatives is critical in bioinformatics
  - Recommend setting threshold of detection based on number of markers with captures[6]
- Limiting for confirming detection of rare species or sampling in novel environments

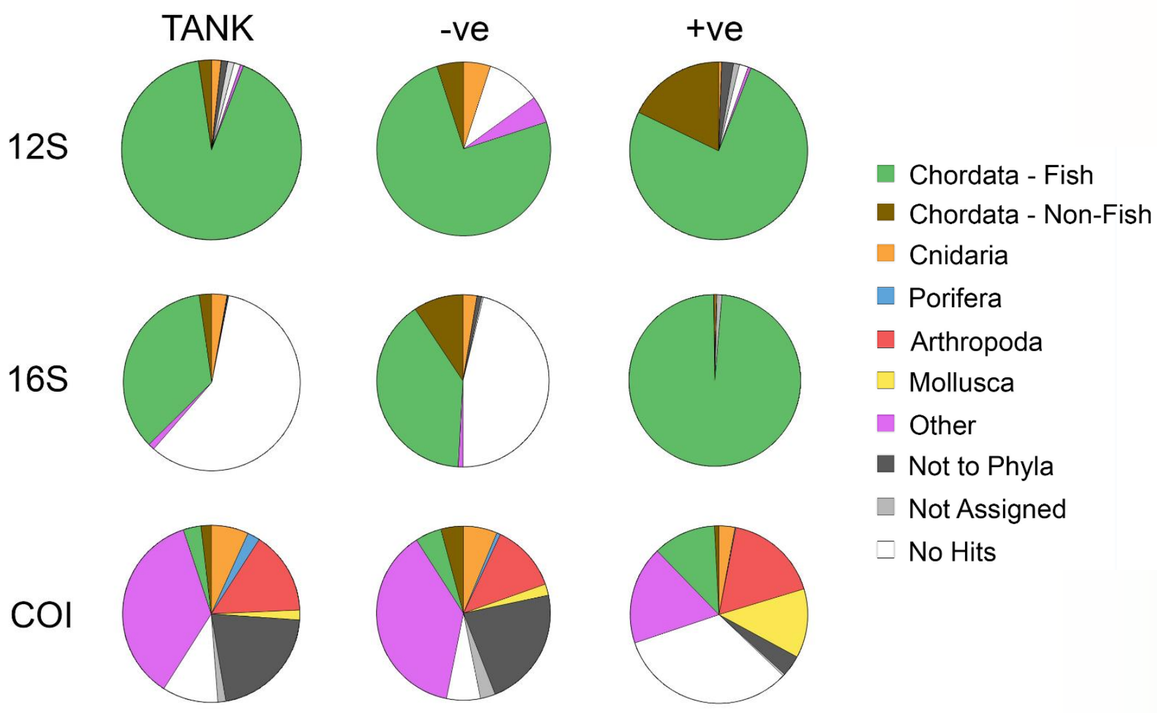

Figure 2. Proportion of OTUs that were generated for each marker, with proportions of OTUs used as a proxy for proportion of species identifications made. Chordata was additionally broken down into Fish and Non-Fish taxa for specificity.

STANDARDIZATION OF BIOINFORMATIC PROCESSING

- Platforms, protocols, and parameters for bioinformatic processing is highly variable and under-described in the literature
  - Limits reproducibility and comparability across studies
  - Adoption of publicly accessible analysis platforms like mBRAVE will improve reproducibility and minimize variability

WHERE TO GO FROM HERE?

- Need to develop better understand of how increasing biodiversity complicates recovery of species ID
- Adoption of multi-marker approaches with a standardized bioinformatics pipeline
  - Adoption of PCR-free sequencing to eliminate marker biases?

REFERENCES

1. Kelly, R.P., Port, J.A., Yamahara, K.M., and Crowder, L.B. (2014). *PLoS One*. **9**(1), e86175.
2. Deagle *et al.* (2009). *Mol. Ecol.,* **18**, 2022-2038.
3. Miya *et al.* (2015). *R. Soc. Open Sci.,* **2,** 150088.
4. Leray *et al.* (2013). *Front. Zool,* **10,** 1-14.
5. Pitz *et al.* (2017). *Proc. Of TDWG*, **1**, e20548.
6. Evans *et al.* (2016). *Can. J. Fish. Aquat. Sci.* **74,** 1362-1374.

ACKNOWLEDGEMENTS

The authors would like to thank the staff at Ripley's Aquarium of Canada, particularly Meredith Schneider and Leah Neal, for assisting with sample collection and access to the sampling site. The authors would like to thank the staff at the Advanced Analysis Center – Genomics Facility at the University of Guelph for assisting with troubleshooting library preparation and conducting the next-generation sequencing runs.

FUNDING SOURCES